

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶: C12N 15/82, C07K 14/415, C12N 1/21, A01N 65/00, A01H 5/00	A1	(11) International Publication Number: WO 95/11306 (43) International Publication Date: 27 April 1995 (27.04.95)
(21) International Application Number: PCT/EP94/03449 (22) International Filing Date: 20 October 1994 (20.10.94) (30) Priority Data: 9321714.9 21 October 1993 (21.10.93) GB (71) Applicant (for all designated States except AT DE US): SANDOZ LTD. [CH/CH]; Lichtstrasse 35, CH-4002 Basle (CH). (71) Applicant (for DE only): SANDOZ-PATENT-GMBH [DE/DE]; Humboldtstrasse 3, D-79539 Lörrach (DE). (71) Applicant (for AT only): SANDOZ-ERFINDUNGEN VERWALTUNGSGESELLSCHAFT MBH [AT/AT]; Brunner Strasse 59, A-1230 Vienna (AT). (72) Inventors; and (75) Inventors/Applicants (for US only): KRAGH, Karsten, Matthias [DK/DK]; Stavtrupvej 139 A, DK-8260 Viby J (DK). MIKKELSEN, Jørn, Dalgaard [DK/DK]; Engkaer 38, DK-2650 Hvidovre (DK). NIELSEN, Klaus, Kristian [DK/DK]; Nebbegardsbåkken 9, DK-2400 Copenhagen NV (DK).	(74) Common Representative: SANDOZ LTD.; Patents & Trade-marks Div., Lichtstrasse 35, CH-4002 Basle (CH). (81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ). Published <i>With international search report.</i>	
(54) Title: ANTI-MICROBIAL PROTEINS (57) Abstract An anti-microbial protein comprising a peptide having the amino acid sequence: AA ₁ -AA ₂ -AA ₃ -Cys-AA ₅ -AA ₆ -AA ₇ -AA ₈ -AA ₉ -Cys-AA ₁₁ -AA ₁₂ -AA ₁₃ -AA ₁₄ -Cys-Cys-AA ₁₇ -AA ₁₈ -AA ₁₉ -AA ₂₀ -AA ₂₁ -Cys-AA ₂₃ -AA ₂₄ -AA ₂₅ -AA ₂₆ -AA ₂₇ -AA ₂₈ -Cys-AA ₃₀ . An anti-microbial protein having the sequence depicted in any one of SEQ ID Nos. 1-3. Recombinant DNA encoding such proteins. A vector comprising such DNA which is expressible in plants and which is linked to a plant operable promoter and terminator. Plants transformed with such recombinant DNA; the progeny of such plants which contain the DNA stably incorporated and heritable in a Mendelian manner, and/or the seeds of such plants or such progeny.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LJ	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

ANTI-MICROBIAL PROTEINS

The present invention relates to anti-microbial proteins isolatable from sugar beet.

According to the present invention there is provided anti-microbial protein comprising a peptide having the amino acid sequence: AA₁-AA₂-AA₃-Cys-AA₅-AA₆-AA₇-AA₈-AA₉-Cys-AA₁₁-AA₁₂-AA₁₃-AA₁₄-Cys-Cys-AA₁₇-AA₁₈-AA₁₉-AA₂₀-AA₂₁-Cys-AA₂₃-AA₂₄-AA₂₅-AA₂₆-AA₂₇-AA₂₈-Cys-AA₃₀, wherein "AA" designates any one of the commonly found 20 amino acids. It is preferred that AA₇ is tyr; AA₁₄ is tyr; and AA₁₈ is lys. In addition it is still more preferred that AA₂₄ is val; AA₂₆ is arg; and AA₂₇ is ala.

An anti-microbial protein includes a protein (alone or in combination with another material) which is toxic or growth inhibitory under any circumstances to any micro-organism, including bacteria, (most particularly Gram positive bacteria), viruses and particularly fungi. Such anti-microbial proteins include those that exhibit anti-microbial activity upon contact with a micro-organism and those that are anti-microbial as a consequence of assimilation or respiration thereof.

The invention also includes an anti-microbial protein having the sequence depicted in any one of SEQ ID Nos. 1-3.

The invention still further includes pure protein which is substantially similar to any one of the above mentioned proteins.

By "substantially similar" is meant pure proteins having an amino acid sequence which is at least 85% similar to the sequence of the proteins according to the invention. It is preferred that the degree of similarity is at least 90%, and still more preferred that the degree of similarity is at least 95%.

In the context of the present invention, two amino acid sequences with at least 85%, 90% or 95% similarity to each other have at least 85%, 90%, or 95% identical or conservatively replaced amino acid residues in a like position when aligned optimally allowing for up to 2

gaps with the *proviso* that in respect of each gap a total not more than 3 amino acid residues is affected. In the case of the proteins specifically depicted in SEQ ID Nos. 1 and 2, the number of gaps may be increased to 4 with the *proviso* that in respect of each gap a total of not more than 5 amino acid residues is affected.

For the purpose of the present invention conservative replacements may be made between amino acids within the following groups:

- (i) Serine and Threonine;
- (ii) Glutamic acid and Aspartic acid;
- (iii) Arginine and Lysine;
- (iv) Asparagine and Glutamine;
- (v) Isoleucine, Leucine, Valine and Methionine;
- (vi) Phenylalanine, Tyrosine and Tryptophan
- (vii) Alanine and Glycine

The invention still further includes pure proteins which are at least 90% identical to the anti-microbial proteins according to the invention, as well as pure proteins which have at least 90% of the specific activity thereof. For the purposes of the present application, specific activity is a measurement of the amount of growth or replication inhibition produced by a specified quantity of the protein on a specified quantity of a specified micro-organism.

The invention still further includes said pure proteins in combination with at least one protein selected from the group consisting of those depicted in SEQ ID Nos. 4-6. Such combined proteins may be further combined with one or more of the known "pathogenesis-related proteins". Infection of plants with fungal or viral pathogens may induce a systemic synthesis of about 10 families of homologous pathogenesis-related proteins (PR proteins) in vegetative tissues. Such PR-proteins have been classified into 5 groups. The PR-2, PR-3 and PR-5 proteins are beta-1,3-glucanase, chitinases and thaumatin-like proteins respectively. Specific functions have not been assigned to the PR-1 and PR-4 groups of proteins. The PR-4 proteins are similar to C-terminal domains of prohevein and the putative wound-induced WIN proteins of potato, thus lacking the N-terminal hevein domain. It is particularly preferred that the proteins according to the invention are combined with one or more proteins which are the

basic counter parts of the P-R 4 group of proteins, meaning the basic counter part of proteins similar to the C-terminal domains of prohevein and the putative wound-induced WIN proteins of potato. It is particularly preferred that the basic counter-part of the said pathogenesis-related proteins is a chitin-binding WIN protein, in particular that produced by barley grain or stressed barley leaves.

The invention still further includes recombinant DNA comprising a sequence encoding a protein having the amino acid sequence of the above disclosed anti-microbial proteins. In particular the DNA may encode at least one of the proteins the sequences of which are depicted in SEQ ID Nos. 1-3, optionally in addition to at least one of the proteins the sequences of which are depicted in SEQ ID Nos. 4-6. The recombinant DNA may further encode a protein having herbicide resistance, plant growth-promoting, anti-fungal, anti bacterial, anti-viral and/or anti-nematode properties. In the case that the DNA is to be introduced into a heterologous organism it may be modified to remove known mRNA instability motifs (such as AT-rich regions) and polyadenylation signals (if any are present), and/or codons which are preferred by the organism into which the recombinant DNA is to be inserted may be used so that expression of the thus modified DNA in the said organism yields substantially similar protein to that obtained by expression of the unmodified recombinant DNA in the organism in which the anti-microbial protein according to the invention is endogenous.

The invention still further includes recombinant DNA which is "similar" to that mentioned above. By "similar DNA" is meant a sequence which is complementary to a test sequence which is capable of hybridizing to the inventive recombinant sequence. When the test and inventive sequences are double stranded the nucleic acid constituting the test sequence preferably has a T_m within 20°C of that of the inventive sequence. In the case that the test and inventive sequences are mixed together and denatured simultaneously, the T_m values of the sequences are preferably within 10°C of each other. More preferably the hybridization is performed under stringent conditions, with either the test or inventive DNA preferably being supported. Thus either a denatured test or inventive sequence is preferably first bound to a support and hybridization is effected for a specified period of time at a temperature of between 50 and 70°C in double strength citrate buffered saline (SSC) containing 0.1% SDS

followed by rinsing of the support at the same temperature but with a buffer having a reduced SSC concentration. Depending upon the degree of stringency required, and thus the degree of similarity of the sequences, such reduced concentration buffers are typically single strength SSC containing 0.1%SDS, half strength SSC containing 0.1%SDS and one tenth strength SSC containing 0.1%SDS. Sequences having the greatest degree of similarity are those the hybridization of which is least affected by washing in buffers of reduced concentration. It is most preferred that the test and inventive sequences are so similar that the hybridization between them is substantially unaffected by washing or incubation in one tenth strength sodium citrate buffer containing 0.1%SDS.

The invention still further includes a DNA sequence which is complementary to one which hybridizes under stringent conditions with the recombinant DNA according to the invention.

Also included in the present invention is: a vector which contains the above disclosed DNA which is expressible in plants and linked to a plant operable promoter and terminator; plants transformed with such DNA; the progeny of such plants which contain the DNA stably incorporated and heritable in a Mendelian manner, and/or the seeds of such plants and such progeny. The transformed plants are made by known methods and include regeneration of plant cells or protoplasts transformed with the DNA of the invention according to a variety of known methods (Agrobacterium Ti and Ri plasmids, electroporation, micro-injection, micro-projectile gun etc). The transformed cell may in suitable cases be regenerated into whole plants in which the nuclear material is stably incorporated into the genome. Both monocot and dicot plants may be obtained in this way. Examples of transformed plants according to the present invention include: fruits, including tomatoes, mangoes, peaches, apples, pears, strawberries, bananas, and melons; field crops such as canola, sunflower, tobacco, sugar beet, small grain cereals such as wheat, barley and rice, maize and cotton, and vegetables such as potato, carrot, lettuce, cabbage and onion. The preferred plants are sugar beet and maize.

The invention still further includes protein derived from expression of the said DNA, and anti-microbial protein produced by expression of the recombinant DNA within plants transformed therewith.

The invention still further includes an anti-microbial composition containing one or more of the proteins according to the invention; a process for combatting fungi which comprises exposing them to such proteins, and an extraction process for obtaining anti-microbial proteins from organic material containing them comprising submitting the material - preferably in the form of a micro-organism - to maceration and solvent extraction. It will be appreciated that the anti-microbial protein exhibits little, if any, anti-microbial effect on the micro-organism which is the source of the organic material referred to in the previous sentence.

The invention will be further apparent from the following description and the associated drawings and sequence listings.

Figure 1 shows a typical elution profile of intercellular washing fluid from a CM-Sephacrose column;

Figure 2 shows a typical elution profile from a Mono S FPLC column of the 0.3M NaCl fraction shown in Figure 1;

Figure 3 shows a typical elution profile from an RP-HPLC column of the proteins represented by peak 3 in Figure 2;

Figure 4 shows the anti-fungal activity of 10ug of the protein represented by peaks 3.1 and 3.2 in Figure 3 (SEQ ID Nos. 1 and 2 respectively);

Figure 5 shows a Sephadex G75 gel filtration chromatogram of the 0.1M fraction depicted in Figure 1;

Figure 6 shows a typical elution profile from a Mono S FPLC column of peak (pool) V depicted in Figure 5;

Figure 7 shows the anti-fungal activity of 10ug of the protein depicted in peaks 1 and 2 (SEQ ID No. 3) in Figure 6.

Figure 8A shows the combined anti-fungal activity of 2ug of each the proteins represented by SEQ ID Nos 2 and 3; Figure 8B shows the combined anti-microbial activity of 2ug of each of the proteins represented by SEQ ID Nos 2 and 5.

Figures 9A and 9B show the morphology of hyphae of *C. beticola* grown in the absence of the anti-microbial proteins of the invention. Figures 9C and 9D show the hyphae when the fungus is grown for 48 hours in the presence of 2ug of the proteins having the sequences depicted in SEQ ID Nos 2 and 3 respectively. The assay is performed in micro-

titre plates as indicated below; magnification Figure 9A, 76x; Figures 9B-9D, 294x.

SEQ ID No. 1 shows the amino acid sequence of protein represented by peak 3.1 in Figure 3; SEQ ID No. 2 shows the amino acid sequence of protein represented by peak 3.2 in Figure 3; and SEQ ID No. 3 shows the amino acid sequence of protein represented by peaks 1 and 2 in Figure 6; SEQ ID Nos. 4-6 show the amino acid sequences of three known anti-fungal proteins; SEQ ID No. 7 shows the nucleotide sequence of the cDNA encoding the protein depicted in SEQ ID No. 3; and SEQ ID No. 8 shows the translation product of the transcript encoded by the cDNA sequence depicted in SEQ ID No. 7, which product includes N- and C- terminal extensions of the protein depicted in SEQ ID No. 3.

Isolation of Intercellular Washing Fluid

IWF is isolated from 500-700 gram sugar beet leaves by submerging them in 20mM HAc (pH 4.5). The thus submerged leaves are then placed in an exicator and vacuum infiltrated for 5 min at 4 torr (max). Following air-drying of the leaf surface, the IWF is collected by centrifugation at 500g for 15 min in 500 ml centrifuge tubes.

Cation exchange chromatography

The thus obtained IWF is fractionated by cation exchange chromatography on a 10 ml CM-Sepharose column (Pharmacia LKB) pre-equilibrated in starting buffer (20mM HAc (pH 4.5)). The fractionation is performed at 4°C at a flow rate of 25 ml/h. Fractions of 3 ml are collected. Proteins not bound to the column are removed by extensive washing of the column with starting buffer. Bound proteins are eluted by applying to the column further starting buffer comprising stepwise increased salt concentrations: viz, 0.1 M NaCl, 0.3 M NaCl and 0.5 M NaCl. The absorbance at 280nm of the eluate is measured, and fractions judged to comprise protein are tested for their anti-fungal activity against *C. beticola* using the microtiter plate bioassay described previously (PCT Patent Application No. PCT/DK92/00108, Publication No. WO 92/17591, now assigned to Sandoz LTD).

A typical elution profile is shown in Fig. 1. The eluates resulting from application to the column of starting buffer comprising the 0.3M NaCl and 0.1M NaCl are further purified as described below.

Purification of antifungal proteins in the 0.3 M NaCl eluate from the CM-Sephacrose.**FPLC Chromatography**

The 0.3M NaCl protein fraction is desalted by overnight dialysis (MW cut off: 3 kDa) against 20mM HAc (pH 4.5) at 4°C. Betaine is added at a concentration of 5% (w/v) to the thus dialysed protein fraction. Four ml of the resulting solution is then fractionated by cation exchange fast protein liquid chromatography (FPLC) using a Mono S HR 5/5 column (Pharmacia LKB) equilibrated in 20 mM HAc (pH 4.5) containing 5% (w/v) betaine (A-buffer). Bound proteins are eluted with a linear salt gradient from 0 to 0.3 M NaCl in 30 ml of the A-buffer followed by a step elution with 1.0 M NaCl in the same buffer. Flow rate is 1 ml/min.

Figure 2 shows that the 0.3M NaCl fraction contains a number of distinct proteins, the quantitatively most significant of which are designated as peaks 1-5. When separated by SDS-PAGE using the Phast System (Pharmacia LKB), silver stained 10-15% gradient Phast gels or High Density gels (Pharmacia LKB) reveal that each of the peaks 1-5 contains 2-5 protein bands.

Reverse Phase HPLC

Protein peak 3 (depicted in Figure 2) from the Mono S column is further purified by reverse phase (RP-) HPLC on a Vydac C₄ silica column (The Separations Group, CA, USA). The solvent system is A: 0.1% TFA in water and B: 0.1% TFA in acetonitrile. Proteins are eluted with a linear gradient of 5 to 45% of the B-buffer applied in 18 min after sample loading followed by 60% B-buffer in 2 min. Flow rate is 0.7 ml/min. Protein is detected by monitoring the absorbance of the eluate at 214 and 280nm. Discrete protein peaks are collected and lyophilized. The thus lyophilized proteins are washed twice with water, re-lyophilized and subsequently resolved in 10mM Tris-HCl (pH 8.0), prior to analysis of purity and anti-fungal activity.

Figure 3 shows that peak 3 from Figure 2 is separated into four peaks (designated as 3.1-3.4 in Figure 3) on the RP column. SDS-PAGE of peaks 3.1-3.4 reveals that each is composed of pure proteins having a molecular weight of about 7 kDa (peaks 3.1, 3.2 and 3.3) and 2.5-3 kD (3.4).

Purification of antifungal proteins in the 0.1 M NaCl eluate from the CM-Sepharose.

Five ml of the 0.1 M NaCl eluate from Figure 1 is fractionated by gel filtration on a 370 ml Sephadex G-75 column (Pharmacia LKB) equilibrated in 50 mM MES (pH 6.0) at a flow rate of 20 ml/hour. Fractions of 10 ml are collected. Such fractions are subsequently pooled in six larger fractions, I-VI, containing approximately 50 ml each (Fig. 5).

Cation exchange (Mono S)

Pool V (shown in Figure 5) from the Sephadex G-75 column is loaded onto a Mono S FPLC column equilibrated in buffer A: 50 mM MES (pH 6.0), containing 5% betaine (w/v). After washing with the A-buffer, bound proteins are eluted at a flow rate of 1ml/min with a linear gradient from 0 to 0.5 M NaCl in 30 ml of the A-buffer (Figure 6). When protein represented by peaks 1 and 2 is subjected to SDS gel electrophoresis in the presence of DTT two distinct bands are observed. The first band has a molecular weight of between 2.5 and 3kDa, and the second band, which represents an unreduced dimer of the protein of the first band, has a molecular weight of about 5kDa.

IDENTIFICATION OF ANTIFUNGAL PROTEINS**Antifungal activity**

Figures 4, and 7-9 show that the proteins represented by peaks 3.1 and 3.2 (SEQ ID Nos 1 and 2 respectively) in Figure 3 and peaks 1 and 2 (SED ID No 3) in Figure 6 have significant anti-fungal activity, *inter alia*, against *C. beticola*. The inhibition of fungal growth is measured in 96 well micro-titre plates at 620nm, essentially as described in WO 92/17591. Figure 9 shows that *C. beticola* treated with such protein possess stunted hyphae which are thickened and more branched (Fig 9C, 9D) when compared to those of fungus grown in the absence of the anti-microbial proteins (Figures 9A, 9B).

The proteins, either alone or in combination with WIN N (which is purified from barley grain or stressed barley leaf as described by Hejgaard *et al* (FEBS Letters, 307, 389-392 (1992)), and/or a protein having a sequence corresponding to at least one of those given in SEQ ID Nos. 4-6, are incubated with spores of *C. beticola*. The assay mix (240ul) contains 100ul of potato dextrose broth (Difco), 40ul protein sample (or buffer control) in 100mM Tris and 20mM NaCl (pH 8.0) as well as approximately 400 spores in 100ul water. The micro-titre

plates are sealed with tape to avoid evaporation and contamination and subsequently incubated at room temperature on an agitator operated at 200 rpm. The absorbance at 620nm is measured each day for 8 days and plotted for each concentration of protein vs time.

Amino acid sequencing

The purified antifungal proteins corresponding to peaks 3.1 and 3.2 in Figure 3 (SEQ ID Nos 1 and 2 respectively) which originate from the 0.3 M NaCl eluate from the CM-Sepharose column, (Figure 1); and protein corresponding to peaks 1 and 2 in Figure 6 (SEQ ID No 3) which originates from the 0.1M NaCl eluate from the said CM-Sepharose column are carboxymethylated and subjected to RP-HPLC on a Vydac C₄ column. The solvent system is A: 0.1% TFA in water and B: 0.1% TFA in acetonitrile. The proteins elute as single peaks with slightly different retention times. The C-terminal sequences of the proteins is obtained by cleavage thereof with *endo*-R-proteinase and subsequent purification by RP-HPLC on a Vydac C₁₈ column.

Production of transformed plants

The genes encoding proteins according to the invention are introduced into plants. Based on gene specific primers, the coding regions of the genes encoding the proteins are synthesized from corresponding mRNA using PCR. After addition of promoter and terminator sequences, the genes encoding the said proteins are introduced into a plant transformation vector. The vector may optionally include a gene encoding a WIN protein, such as that obtained from stressed barley leaf or barley grain, and/or a gene encoding a protein depicted in SEQ ID No. 4, SEQ ID No. 5 and/or SEQ ID No. 6, and/or a gene encoding a chitinase and or a glucanase. The preferred chitinase is the chitinase 4 described in PCT Patent Application No. PCT/DK92/00108 (Publication No. WO 92/17591). *Agrobacterium tumefaciens*, for example, may be transformed with these vectors. Plant cells are then treated with such transformed *Agrobacterium*, and the thus transformed plant cells are regenerated into whole plants, in which the new nuclear material is stably incorporated into the genome. It will be appreciated, however, that the DNA encoding a protein according to the present invention, (or combination of such proteins), optionally further encoding other proteins, may be introduced into plant cells by other known methods, including use of a micro-projectile gun, electroporation, electro-transformation, and micro-injection etc, and that regeneration of

transformed plant cells is carried out according to methods known to the skilled man, including treatment of the cells with cytokines where this is necessary or desirable in order to improve the regeneration frequency.

Moreover, suitable micro-organisms (i.e. those in which the production of the present inventive proteins is not substantially toxic) may be transformed with a vector comprising the gene (or genes) encoding the protein so that the transformed micro-organisms produce such protein. The micro-organisms may further comprise the genes encoding other proteins, such as a WIN protein and/or a protein the sequence of which is depicted in one or more of SEQ IDs No. 4 - 6. Furthermore, such other proteins may further comprise various chitinases and/or glucanases. A particularly preferred such other protein is the chitinase 4 as described in PCT Patent Application No. PCT/DK92/00108 (Publication No. WO 92/17591).

These micro-organisms may then be used to combat plant pathogens. For example the transformed micro-organisms may be dried and sprayed onto infected plants or plants at risk of infection.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Sandoz Ltd
 (B) STREET: Lichtstrasse 35
 (C) CITY: Basel
 (D) STATE: BS
 (E) COUNTRY: Switzerland
 (F) POSTAL CODE (ZIP): CH-4002
 (G) TELEPHONE: 061-324-1111
 (H) TELEFAX: 061-322-7532
 (I) TELEX: 965-05055

(A) NAME: Sandoz Erfindungen Verwaltungsgesellschaft
 MBH

(B) STREET: Brunnerstrasse 59
 (C) CITY: Vienna
 (E) COUNTRY: Austria
 (F) POSTAL CODE (ZIP): A-1230

(A) NAME: Sandoz-Patent GMBH
 (B) STREET: Humboldtstrasse 3
 (C) CITY: Loerrach
 (E) COUNTRY: Germany
 (F) POSTAL CODE (ZIP): D-7850

(ii) TITLE OF INVENTION: Anti-microbial proteins

(iii) NUMBER OF SEQUENCES: 8

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
 (B) COMPUTER: IBM PC compatible
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 91 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Beta vulgaris

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ile	Thr	Cys	Gly	Leu	Val	Ala	Ser	Lys	Leu	Ala	Pro	Cys	Ile	Gly	Tyr
1				5				10					15		
Leu	Gln	Gly	Ala	Pro	Gly	Pro	Ser	Ala	Gly	Cys	Cys	Gly	Gly	Ile	Lys
			20					25					30		

Gly Leu Asn Ser Ala Ala Ala Ser Pro Ala Asp Arg Lys Thr Ala Cys
 35 40 45
 Thr Cys Leu Lys Ser Ala Ala Thr Ser Ile Lys Gly Ile Asn Tyr Gly
 50 55 60
 Lys Ala Ala Ser Leu Pro Arg Gln Cys Gly Val Ser Val Pro Tyr Ala
 65 70 75 80
 Ile Ser Pro Asn Thr Asn Cys Asn Ala Ile His
 85 90

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 91 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Beta vulgaris
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ile Thr Cys Gly Leu Val Ala Ser Lys Leu Ala Pro Cys Ile Gly Tyr
 1 5 10 15
 Leu Gln Gly Ala Pro Gly Pro Ser Ala Gly Cys Cys Gly Gly Ile Lys
 20 25 30
 Gly Leu Asn Ser Ala Ala Ala Ser Pro Ala Asp Arg Lys Thr Ala Cys
 35 40 45
 Thr Cys Leu Lys Ser Ala Ala Thr Ser Met Lys Gly Ile Asn Tyr Gly
 50 55 60
 Lys Ala Ala Ser Leu Pro Arg Gln Cys Gly Val Ser Ile Pro Tyr Ala
 65 70 75 80
 Ile Ser Pro Asn Thr Asn Cys Asn Ala Ile His
 85 90

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Beta vulgaris

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ser Gly Glu Cys Asn Met Tyr Gly Arg Cys Pro Pro Gly Tyr Cys Cys
 1 5 10 15
 Ser Lys Phe Gly Tyr Cys Gly Val Gly Arg Ala Tyr Cys Gly
 20 25 30

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 46 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Beta vulgaris

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ala Ile Cys Lys Lys Pro Ser Lys Phe Phe Lys Gly Ala Cys Gly Arg
 1 5 10 15
 Asp Ala Asp Cys Glu Lys Ala Cys Asp Gln Glu Asn Trp Pro Gly Gly
 20 25 30
 Val Cys Val Pro Phe Leu Arg Cys Glu Cys Gln Arg Ser Cys
 35 40 45

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 46 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Beta vulgaris

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ala Thr Cys Arg Lys Pro Ser Met Tyr Phe Ser Gly Ala Cys Phe Ser
 1 5 10 15
 Asp Thr Asn Cys Gln Lys Ala Cys Asn Arg Glu Asp Trp Pro Asn Gly
 20 25 30
 Lys Cys Leu Val Gly Phe Lys Cys Glu Cys Gln Arg Pro Cys
 35 40 45

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: Beta vulgaris
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
- | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
| Arg | Cys | Ile | Pro | Cys | Gly | Gln | Asp | Cys | Ile | Ser | Ser | Arg | Asn | Cys | Cys | | | | | | | | | | | | | | | |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | | | | | | | | | | | | | | | | |
| Ser | Pro | Cys | Lys | Cys | Asn | Phe | Gly | Pro | Pro | Val | Pro | Arg | Cys | Thr | Asn | | | | | | | | | | | | | | | |
| | | | 20 | | | | 25 | | | | | | 30 | | | | | | | | | | | | | | | | | |

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 550 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: unknown
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: Beta vulgaris
- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 66..293
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:
- | | | | | | | | |
|---|---|------------|------------|------------|------------|--|----|
| ACTCAACAAA | TTCAGAAAAA | AACAGAAGCA | AAAAAAGTTT | ATTGAAAGAG | TAAGTTGAGG | | 60 |
| TGAAA | ATG ATG AAA AGC TTT GTG ATA GTT ATG TTG GTC ATG TCC ATG | | 107 | | | | |
| Met Met Lys Ser Phe Val Ile Val Met Leu Val Met Ser Met | | | | | | | |
| 1 | | 5 | 10 | | | | |
| ATG GTG GCT ACA TCT ATG GCA AGT GGT GAA TGC AAT ATG TAT GGT CGA | | | 155 | | | | |
| Met Val Ala Thr Ser Met Ala Ser Gly Glu Cys Asn Met Tyr Gly Arg | | | | | | | |
| 15 | | 20 | 25 | 30 | | | |
| TGC CCC CCA GGG TAT TGT TGT AGC AAG TTT GGC TAC TGT GGT GTC GGA | | | 203 | | | | |
| Cys Pro Pro Gly Tyr Cys Cys Ser Lys Phe Gly Tyr Cys Gly Val Gly | | | | | | | |
| | 35 | 40 | 45 | | | | |
| CGC GCC TAT TGT GGC GAT GCT GAG CAG AAG GTT GAA GAT CAT CCA TCT | | | 251 | | | | |
| Arg Ala Tyr Cys Gly Asp Ala Glu Gln Lys Val Glu Asp His Pro Ser | | | | | | | |
| | 50 | 55 | 60 | | | | |


```

AAT GAT GCT GAT GTT CCT GAG TTT GTT GGA GCT GGT GCC CCA      293
Asn Asp Ala Asp Val Pro Glu Phe Val Gly Ala Gly Ala Pro
    65              70              75

TGATGCTCGA AGCCAGGTAA TCGTAATGGC ATGGGTTACC TAATAAGTAA ACTCATTGTG      353
CCTAGCTTGC TACATGCTTA TCCACTATAA ATAAGCTCCT ACAGGAGTTG TGTTCCTTCTT      413
TTAATTTTGT AATCAAGGGT TTGACTTTAA TTAATGAGAC CAATGTATAC TTGCATGTCG      473
GATAAATATN TAACTAAGCC ACTCGTATTG GTTTATTATA AACTACTAT AAAAAAAAAA      533
AAAAAAAAAA AAAAAAA      550

```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 76 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Met Met Lys Ser Phe Val Ile Val Met Leu Val Met Ser Met Met Val
 1              5              10              15
Ala Thr Ser Met Ala Ser Gly Glu Cys Asn Met Tyr Gly Arg Cys Pro
    20              25              30
Pro Gly Tyr Cys Cys Ser Lys Phe Gly Tyr Cys Gly Val Gly Arg Ala
    35              40              45
Tyr Cys Gly Asp Ala Glu Gln Lys Val Glu Asp His Pro Ser Asn Asp
    50              55              60
Ala Asp Val Pro Glu Phe Val Gly Ala Gly Ala Pro
    65              70              75

```

Claims

1. An anti-microbial protein comprising a peptide having the amino acid sequence:
AA₁-AA₂-AA₃-Cys-AA₅-AA₆-AA₇-AA₈-AA₉-Cys-AA₁₁-AA₁₂-AA₁₃-AA₁₄-Cys-Cys-AA₁₇-AA₁₈-
AA₁₉-AA₂₀-AA₂₁-Cys-AA₂₃-AA₂₄-AA₂₅-AA₂₆-AA₂₇-AA₂₈-Cys-AA₃₀.
2. An anti-microbial protein according to claim 1, wherein AA₇ is tyr; AA₁₄ is tyr; and AA₁₈ is lys.
3. An anti-microbial protein according to either of claims 1 or 2, wherein AA₂₄ is val; AA₂₆ is arg; and AA₂₇ is ala.
4. An anti-microbial protein having the sequence depicted in any one of SEQ ID Nos 1-3.
5. Pure protein at least 90% similar to the proteins according to any one of claims 1-4.
6. Pure proteins according to any one of claims 1-5, in combination with at least one protein selected from the group consisting of those depicted in SEQ ID Nos. 4 - 6.
7. Pure proteins according to any one of the preceding claims, in combination with at least one further protein having herbicide resistance, plant growth-promoting, anti-fungal, anti-bacterial, anti-viral and/or anti-nematode properties.
8. Recombinant DNA comprising a sequence encoding a protein as claimed in any one of claims 1-5, or recombinant DNA having the sequence depicted in SEQ ID No. 7.
9. A recombinant DNA sequence according to claim 8, which further comprises a sequence encoding at least one of the proteins depicted in SEQ ID Nos. 4 - 6.
10. Recombinant DNA according to either of claims 8 or 9, which further encodes a protein having herbicide resistance, plant growth-promoting, anti-fungal, anti-bacterial, anti-

viral and/or anti-nematode properties.

11. Recombinant DNA according to any one of claims 8 to 10, which is modified in that known mRNA instability motifs or polyadenylation signals are removed and/or codons which are preferred by the organism into which the recombinant DNA is to be inserted are used so that expression of the thus modified DNA in the said organism yields substantially similar protein to that obtained by expression of the unmodified recombinant DNA in the organism in which the protein is endogenous.

12. A DNA sequence which is complementary to one which hybridizes under stringent conditions with the DNA of any one of claims 8 to 11.

13. A vector containing a DNA sequence as claimed in any one of claims 8 to 12, which sequence is expressible in plants and linked to a plant operable promoter and terminator.

14. A biological system which includes DNA as claimed in any one of claims 8 to 12, or the vector of claim 13.

15. Plants transformed with recombinant DNA as claimed in any one of claims 8 to 12, the progeny of such plants which contain the DNA stably incorporated and heritable in a Mendelian manner, and/or the seeds of such plants or such progeny.

16. Protein derived from expression of the DNA as claimed in any one of claims 8 to 12, and anti-microbial protein produced by expression of the recombinant DNA within plants or progeny or seed thereof as claimed in claim 15.

17. An anti-microbial composition containing one or more of the proteins as claimed in any one of claims 1-7 and 16.

18. A process for combatting fungi which comprises exposing them to proteins or compositions as claimed in any one of claims 1-7, 16 and 17.

19. An extraction process for obtaining anti-microbial proteins, as claimed in any one of claims 1-7 or claim 16, from organic material containing them comprising submitting the material to maceration and solvent extraction, characterized in that the material is a micro-organism or a plant.

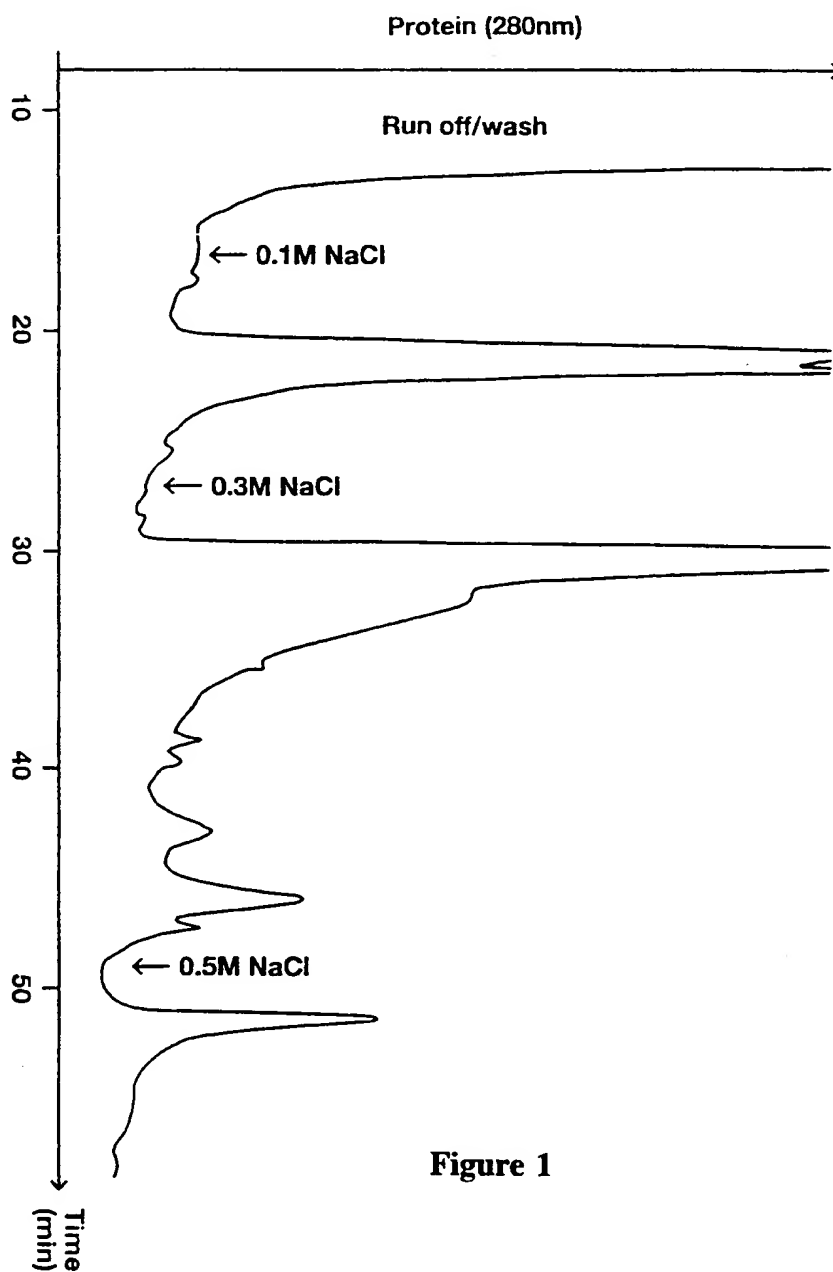
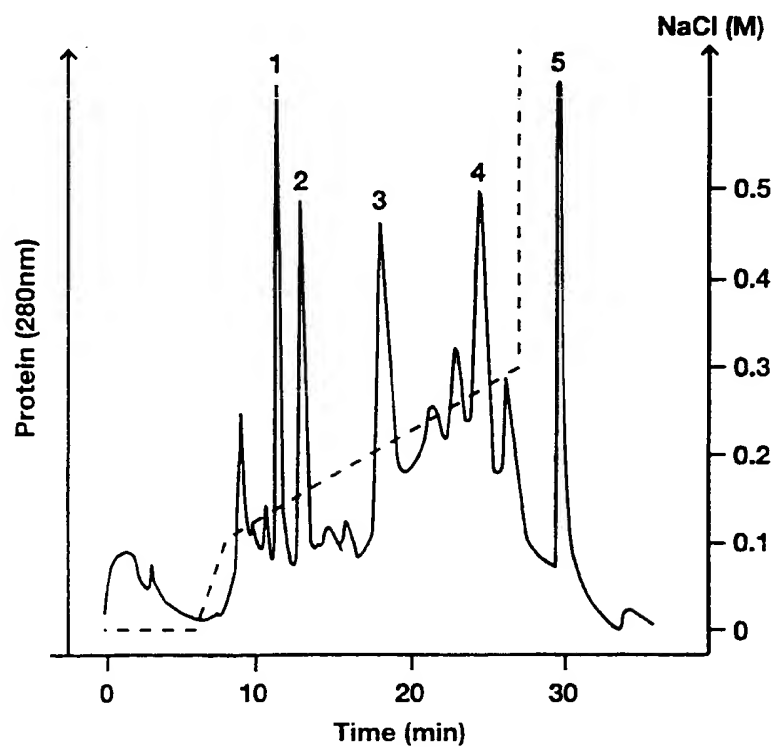


Figure 1

**Figure 2**

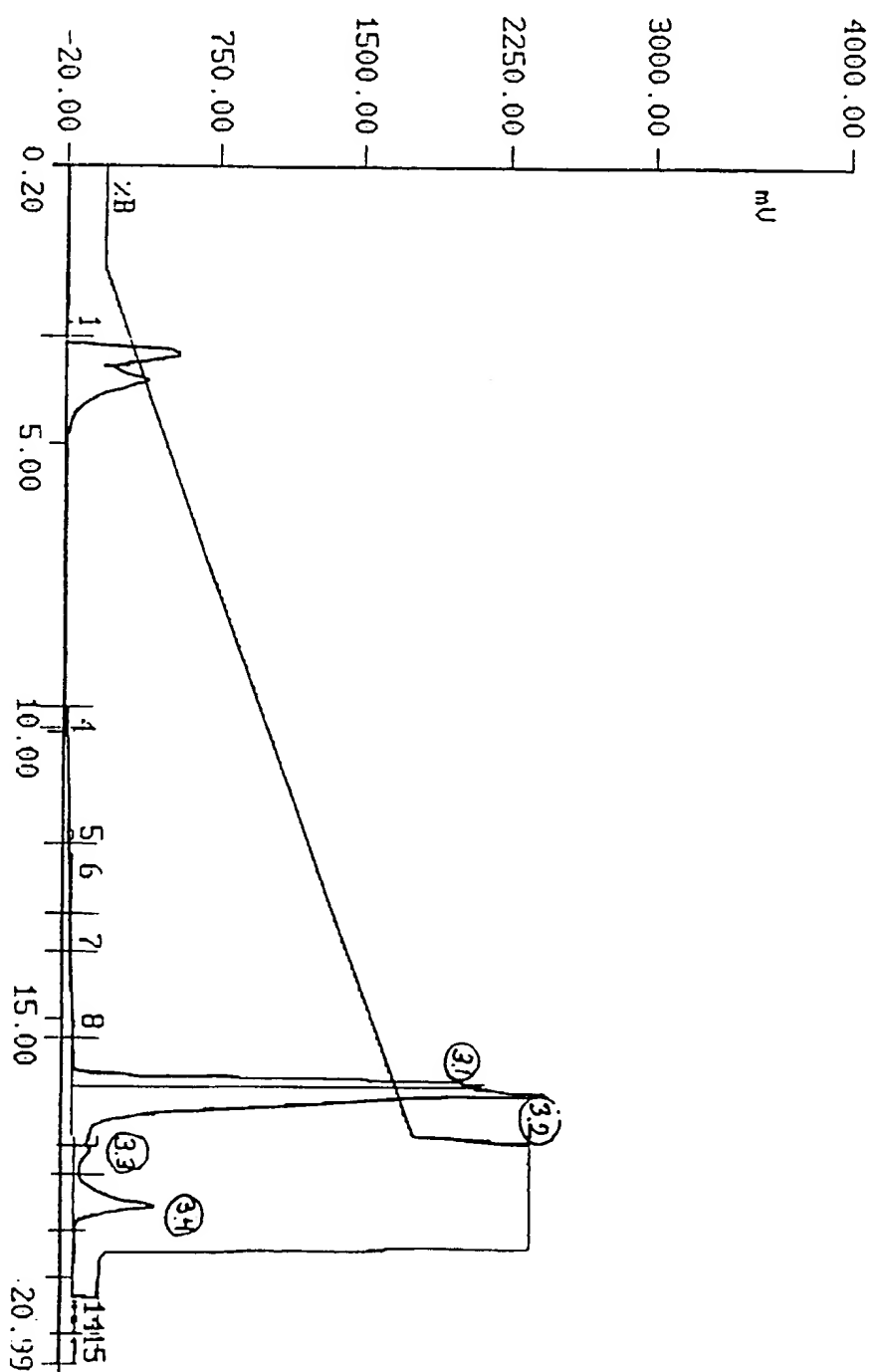


Figure 3

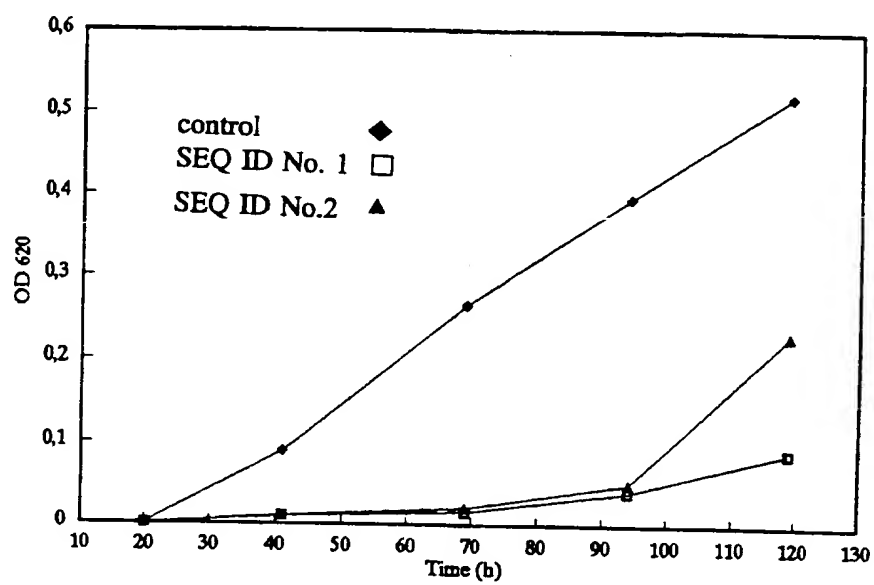
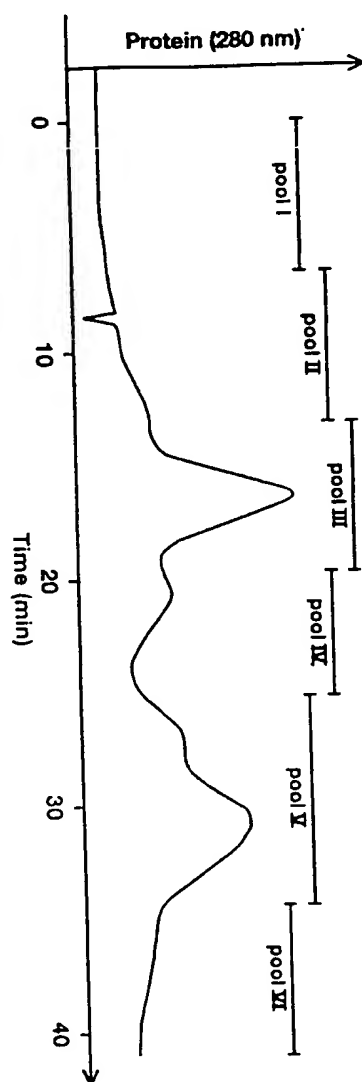


Figure 4

**Figure 5**

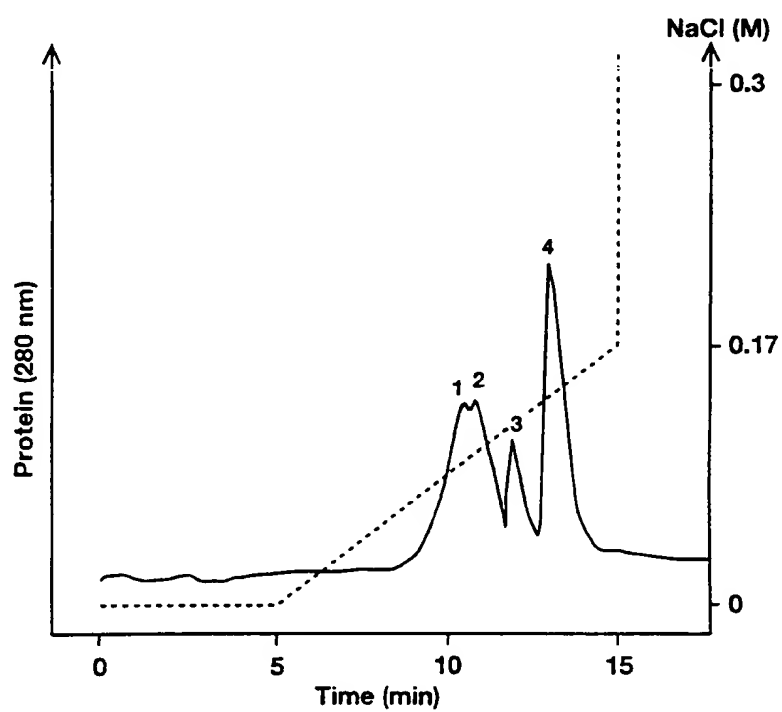


Figure 6

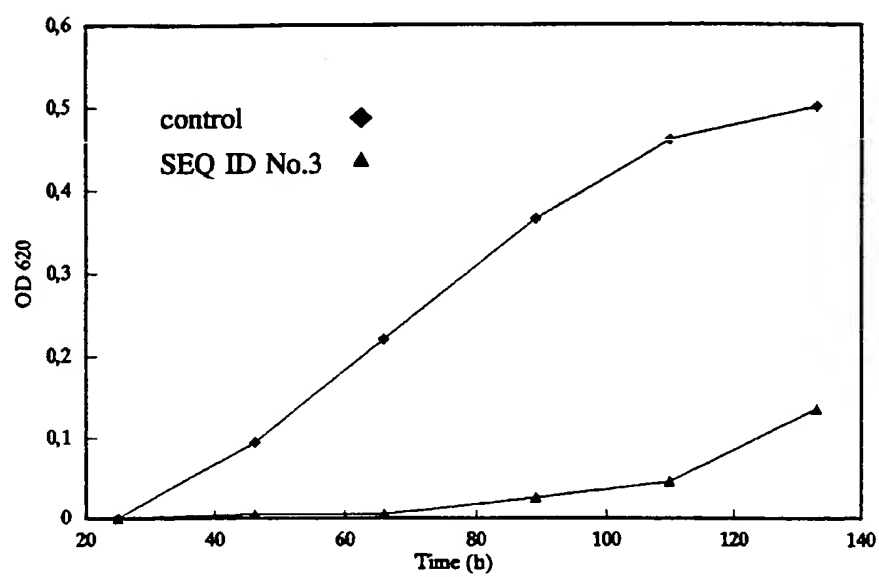


Figure 7

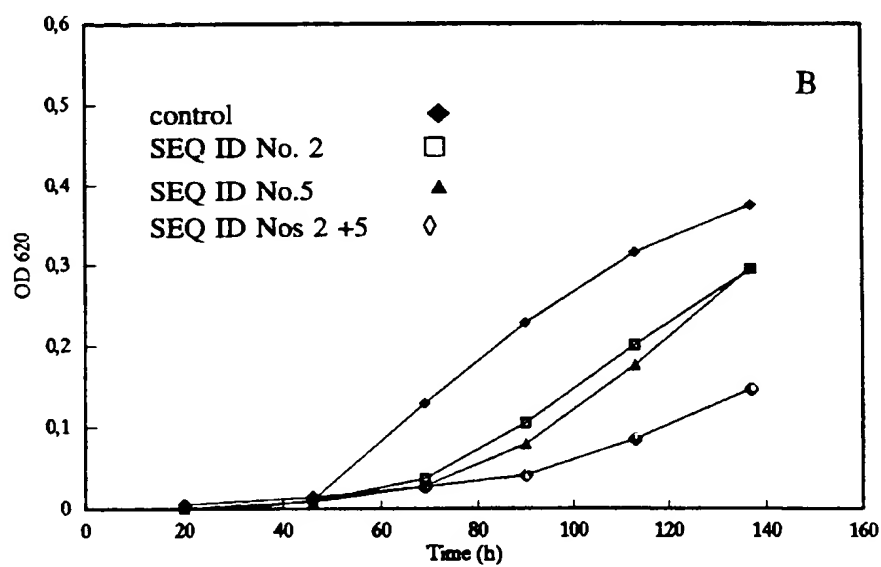
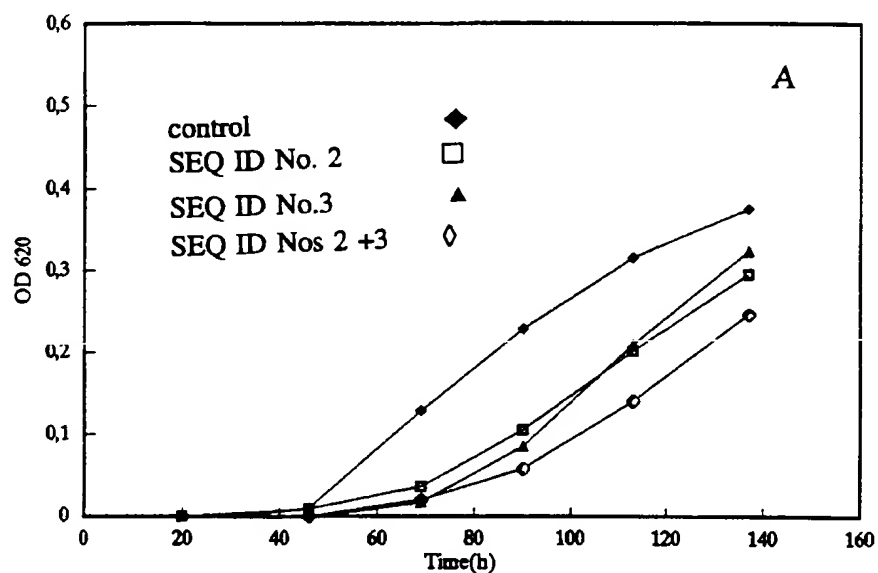


Figure 8

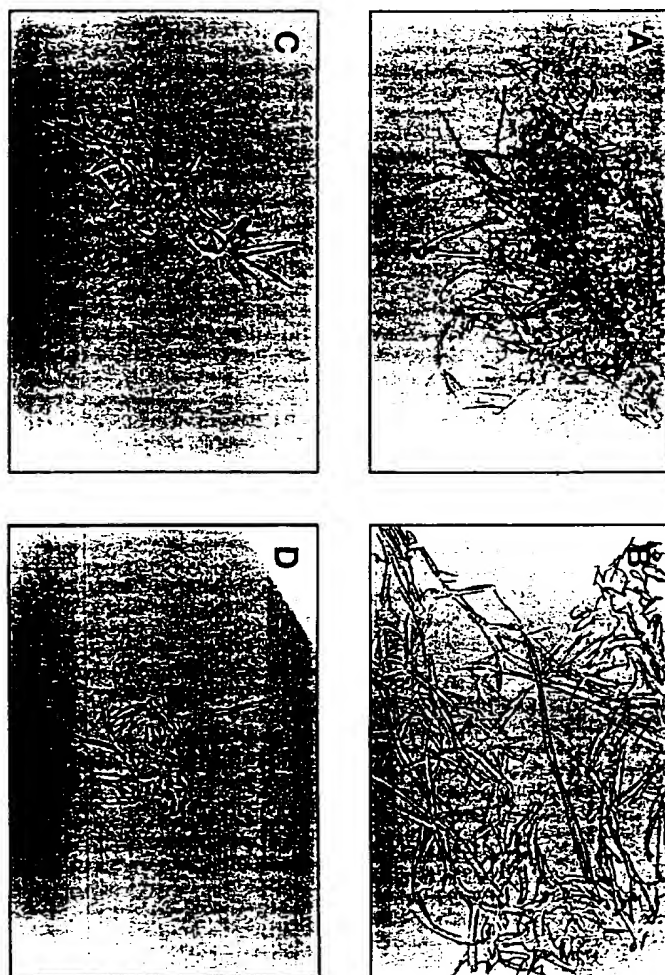


Figure 9

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 94/03449

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/82 C07K14/415 C12N1/21 A01N65/00 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07K C12N A01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO,A,92 21699 (I.C.I PLC.) 10 December 1992 see abstract; claims 1-20; figures 4,8 ---	1,5,8, 10-19
X	WO,A,92 17591 (DANISCO A/S) 15 October 1992	5
Y	see the whole document ---	1,8, 12-19
Y	DE,A,36 42 050 (TAKARA SHUZO CO. LTD.) 11 June 1987 see abstract; claim 1 ---	1,8, 12-19
	--- -/--	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Z" document member of the same patent family

Date of the actual completion of the international search

17 January 1995

Date of mailing of the international search report

06 -02- 1995

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+ 31-70) 340-3016

Authorized officer

Gurdjian, D

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 94/03449

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	PLANT PHYSIOLOGY, vol.91, 1989, BETHESDA US pages 124 - 129 LERNER D.R. ET AL. 'CLONING AND CHARACTERIZATION...' see the whole document -----	1,5,8, 10-19
P,A	WO,A,94 11511 (ZENECA LTD.) 26 May 1994 see the whole document -----	1-19

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 94/03449

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9221699	10-12-92	AU-A- 1780092	08-01-93
		BR-A- 9206105	27-09-94
		EP-A- 0593501	27-04-94
		NZ-A- 243063	27-09-94
WO-A-9217591	15-10-92	AU-A- 1659992	02-11-92
		CZ-A- 9302092	13-04-94
		EP-A- 0579709	26-01-94
		JP-T- 6507070	11-08-94
		NZ-A- 242270	26-07-94
DE-A-3642050	11-06-87	JP-A- 62135492	18-06-87
WO-A-9411511	26-05-94	AU-B- 5340694	08-06-94